Bayesian inference for geometric distribution under a simple step-stress model

Arefi, A.
Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran,
email: arefiahmad@yahoo.com

Razmkhah, M.
Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

In some situations in reliability and survival analysis, the life times of the units in an experiment depend on the number of times the units are switched on and off or the number of shocks they receive. On the other hand, the experiment may not terminate on an adequate time under the normal conditions. This paper proposes a Bayesian inference model for a simple step-stress model with Type-I censored sample. Assuming a cumulative exposure model with lifetimes being geometric distributed, the problems of point and interval estimation of studied in the Bayesian approach. Finally, an example is presented to illustrate the proposed procedure in this paper.

Keywords and phrases: Bayesian confidence interval, Beta prior distribution, Accelerated testing, Order statistics.

AMS Subject Classification 2000: primary 62N05; secondary 62F15.

1 Introduction

There are situations in reliability and survival analysis for which the experiment may not terminate on an adequate time under the normal conditions. In such situations, accelerated life-testing experiments have been offered to obtain adequate life data. See, for example, Nelson (1990) and Bagdonavicius and Nikulin (2002). Step-stress accelerated life testing (SSALT) is a special class of accelerated life-testing for which the stress levels of the experiment change at some pre-specified times. Balakrishnan *et al.* (2009) derived exact inference for simple step-stress model from the exponential distribution when there is time constraint on the duration of the experiment. See also, Balakrishnan and Xie (2007, a) and Balakrishnan and Han (2008). DeGroot and Goel (1979) proposed a Bayesian inference model for SSALT and an criterion optimality for simple SSALT in the framework of Bayesian decision theory. See also, Van Dorp et al. (1996) and Van Dorp and Mazzuchi (2004, 2005) and Erto, and Giorgio (2002). The classical approach treats parameters of life distribution as fixed but unknown constants but Bayes approach considers them as random variables whit a prior distributions for these parameters. Prior distributions are constructed by existing information or subjective judgments.
In some situations, the life times of the units in an experiment depend on the number of times the units are switched on and off or the number of shocks they receive. Let w be the number of switch on and off or shocks the units receive until they fail, so, w is considered as the associated failure time. Here, the life-testing experiment are investigated in a discrete set up. See, Nagaraja (1992) for more details about the results on order statistics of a random sample taken from a discrete population. Censored samples in discrete set up have been studied by some authors. See, for example, Rezaei and Arghami (2002), Davarzani and Parsian (2011) and Balakrishnan *et al.* (2011). This paper proposes a Bayesian inference model for a simple SSALT having only one change between two stress levels, when Type-I censoring is used. We assume that the failure times at each stress level follow a geometric distribution.

The rest of the paper is as follows: In Section 2 some preliminaries are presented. In Section 3, the Byesian estimation of the parameters of the geometric distribution is investigated for a simple step-stress model with Type-I censored sample. Finally, in Section 4, an example is given to illustrate the results of the paper.

2 Preliminaries

Consider a simple step-stress scheme with only two stress levels S_1 and S_2 and assume that the associated distributions at levels S_1 and S_2 are geometric with successive probabilities p_1 and p_2, respectively. The probability mass function (pmf) and cumulative distribution function (cdf) are given by

$$\mathrm{P}_j(X_j = x) = p_j q_j^{x-1}, \quad x = 1, 2, \ldots,$$

$$F_j(x; p_j) = 1 - q_j^x, \quad x = 1, 2, \ldots \tag{1}$$

The parameters of interest in this paper are:(i) the successive probability at level S_j, i.e.,p_j, $(j = 1, 2)$, (ii) the mathematical expectation at level S_j,i.e., $\mu_j = \frac{1}{p_j}$, (iii) the survival function of the level S_j at $x_0 : \bar{F}_j(x_0) = q_j^{x_0}$.
Suppose that the normal conditions (level S_1) of an experiment change to level S_2 at point w_1. Therefore, using (1), the cumulative exposure distribution (ced) $G(x)$ is

$$\begin{aligned} G(x) &= \begin{cases} G_1(x) = F_1(x; p_1), & x = 1, 2, \ldots, w_1, \\ G_2(x) = F_2\big(x - (1 - \log q_1 / \log q_2) w_1; p_2\big), & x = w_1 + 1, w_1 + 2, \ldots, \end{cases} \\ &= \begin{cases} G_1(x) = 1 - q_1^x, & x = 1, 2, \ldots, w_1, \\ G_2(x) = 1 - q_1^{w_1} q_2^{x - w_1}, & x = w_1 + 1, w_1 + 2, \ldots \end{cases} \end{aligned} \tag{2}$$

and the corresponding pmf $g(x)$ is as follows

$$g(x) = \begin{cases} g_1(x) = p_1 q_1^{x-1}, & x = 1, 2, \ldots, w_1, \\ g_2(x) = p_2 q_1^{w_1} q_2^{x-(w_1+1)}, & x = w_1 + 1, w_1 + 2, \ldots. \end{cases}$$

We now introduce some notations will be used throughout the paper: $X_{i:n}$ denotes the ith smallest order statistics in a sample of size n from the geometric distribution. N_1 is the number of observations that are less than or equal to w_1 and N_2 denotes the number of data points that are less than or equal to w_2 and greater than w_1, for which $N_1 + N_2 \leq n$. Using these notation, we will observe the following data set under Type-I censoring scheme:

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{N_1:n} \leq w_1 < X_{N_1+1:n} \leq \cdots \leq X_{N_1+N_2:n} \leq w_2. \tag{3}$$

Notice that in the special case of $N_1 + N_2 = n$, the complete sample is observed. To study the estimation problem of the parameter of interest based on the data set in (3), we need to obtain the joint distribution of $X_{1:n}, \ldots, X_{N_1+N_2:n}$. In order to provide explicit expression for the joint distribution of discrete-order statistics, it is necessary to use the 'tie- run" technique which is defined by Gan and Bain (1995) regarding the number and lengths of runs of tied observations. A subchain $t_{i_1} \leq t_{i_2} \leq \cdots \leq t_{i_n}$ of real numbers is said to have r tie-runs $(1 \leq r \leq n)$ with length z_k $(1 \leq k \leq r)$ for the kth one, if

$$t_{i_1} = \cdots = t_{i_{z_1}} < t_{i_{z_1+1}} = \cdots = t_{i_{z_1+z_2}} < \cdots < t_{i_{n-z_r+1}} = \cdots = t_{i_n},$$

with $\sum_{k=1}^r z_k = n$. Let $X_1, \ldots, X_n$ be iid discrete random variables from the ced in (2). Using the concept of tie-run, given p_1 and p_2, the joint pmf of $X_{1:n}, \cdots, X_{N_1+N_2:n}, N_1, N_2$ is as follows

$$\begin{aligned} L(p_1, p_2) &= \frac{n!}{(n - n_1 - n_2)!} \left(\prod_{j=1}^{r} z_j! \right)^{-1} \prod_{i=1}^{n_1} g_1(x_{i:n}) \\ &\times \prod_{i=n_1+1}^{n_1+n_2} g_2(x_{i:n}) \big(1 - G_2(w_2)\big)^{n-n_1-n_2} \\ &= \frac{n!}{(n - n_1 - n_2)!} \left(\prod_{j=1}^{r} z_j! \right)^{-1} p_1^{n_1} q_1^{d_1} p_2^{n_2} q_2^{d_2}, \end{aligned} \tag{4}$$

where r is equal to the number of tie-runs with length z_k for the jth one, d_1 and d_2 are the observed values of , D_1 and D_2, respectively, where

$$D_1 = \sum_{i=1}^{N_1} X_{i:n} - N_1 + w_1(n - N_1), \tag{5}$$

$$D_2 = \sum_{i=N_1+1}^{N_1+N_2} X_{i:n} - (w_1+1)N_2 + (w_2 - w_1)(n - N_1 - N_2). \tag{6}$$

3 Baysian estimation

In this section we present and illustrate the methodology for obtaining the Bayes estimators.Toward this end, we assume that the parameters p_1 and p_2 behave as independent random variables. Also, suppose the random variable p_j has Beta prior distribution with parameters α_j and β_j $(j = 1, 2)$. That is, the prior density function of $p_j,\ j = 1, 2$, takes the following form

$$\pi_j(p_j) = \frac{1}{\beta(\alpha_j, \beta_j)} p_j^{\alpha_j - 1}(1 - p_j)^{\beta_j - 1}, \quad 0 < p_j < 1. \tag{7}$$

Therefore, by performing some algebraic calculations, it con be shown that, the joint posterior pdf of p_1 and p_2 is

$$\pi(p_1, p_2|data) = \frac{p_1^{a_1-1}(1-p_1)^{b_1-1} p_2^{a_2-1}(1-p_2)^{b_2-1}}{\prod_{j=1}^{2} B(a_j, b_j)}, \tag{8}$$

where $a_j = N_j + \alpha_j$. and $b_j = D_j + \beta_j$ $(j = 1, 2)$, for which D_1 and D_2 are as defined in (5) and (6), respectively. Using (8), the marginal posterior of p_j is

$$\pi_j(p_j|data) = \frac{p_j^{a_j-1}(1-p_j)^{b_j-1}}{B(a_j, b_j)}, \quad j = 1, 2. \tag{9}$$

3.1 Bayesian point estimation

To proceed the problem of Bayes estimation, we use the squared error loss (SEL) function. Let $\hat{\theta}$ be any estimator for θ, then the SEL function is defined by $L(\hat{\theta} - \theta) = (\hat{\theta} - \theta)^2$. Using the SEL function, the Bayes estimate of the unknown parameter is simply the mean of the posterior distribution. It can be shown that the Bayes risk is the variance of the posterior distribution. In the following results, the Bayes estimators for the parameters of interest in this paper are presented. Under the of SEL function, we have

(i) The Bayes estimator for $p_j (j = 1, 2)$ is given by

$$\hat{p}_j = \frac{a_j}{a_j + b_j}; \tag{10}$$

(ii) The Bayes risk associated with $\hat{p}_j$ say $R_{\hat{p}_j}$, is

$$R_{\hat{p}_j} = \frac{a_j b_j}{(a_j + b_j + 1)(a_j + b_j)^2}.$$

the Bayes estimator for $\mu_j,\ j = 1, 2$ is

$$\hat{\mu}_j = \frac{a_j + b_j - 1}{a_j - 1}, \tag{11}$$

Moreover the Bayes risk associated with $\hat{\mu}_j$ say $R_{\hat{\mu}_j}$, is

$$R_{\hat{\mu}_j} = \frac{(a_j + b_j - 2)(a_j + b_j - 1)}{(a_j - 2)(a_j - 1)} - \left(\frac{a_j + b_j - 1}{a_j - 1}\right)^2.$$

For estimating the survival function of the level S_j at x_0, we have (i) The Bayes estimator for $\bar{F}_j(x_0)(j = 1, 2)$ is

$$\hat{\bar{F}}_j(x_0) = \frac{b_j^{[x_0]}}{(a_j + b_j)^{[x_0]}}; \tag{12}$$

where

$$\begin{aligned} b_j^{[x_0]} &= (b_j + x_0 - 1)(b_j + x_0 - 2)\dots(b_j + 1)\ b_j, \\ (a_j + b_j)^{[x_0]} &= (a_j + b_j + x_0 - 1)(a_j + b_j + x_0 - 2)\dots(b_j + a_j), \end{aligned}$$

(ii) The Bayes risk associated with $\hat{\bar{F}}_j(x_0)$ say $R_{\hat{\bar{F}}_j(x_0)}$, is

$$R_{\hat{\bar{F}}_j(x_0)} = \frac{b_j^{[2x_0]}}{(a_j + b_j)^{[2x_0]}} - \left\{\frac{b_j^{[x_0]}}{(a_j + b_j)^{[x_0]}}\right\}^2;$$

where

$$\begin{aligned} b_j^{[2x_0]} &= (b_j + 2x_0 - 1)(b_j + 2x_0 - 2)\dots(b_j + 1)\ b_j, \\ (a_j + b_j)^{[2x_0]} &= (a_j + b_j + 2x_0 - 1)(a_j + b_j + 2x_0 - 2)\dots(b_j + a_j). \end{aligned}$$

3.2 Bayesian interval estimation

Once the posterior probability density function $h(\theta|data)$ of the unknown parameter θ is derived,the interval is $100(1 - \alpha)\%$ Bayesian confidence interval for θ, is

$$P(L \leq \theta \leq U|data) = 1 - \alpha. \tag{13}$$

Using (9) and (13), a $100(1 - \alpha)\%$ Bayesian confidence intervals for $p_j(j = 1, 2)$, say (L_j, U_j), can be derived by solving the following two equations

$$\frac{\alpha}{2} = \int_0^{L_j} \frac{p_j^{a_j-1}(1 - p_j)^{b_j-1}}{B(a_j, b_j)} dp_j, \quad \frac{\alpha}{2} = \int_{U_j}^1 \frac{p_j^{a_j-1}(1 - p_j)^{b_j-1}}{B(a_j, b_j)} dp_j. \tag{14}$$

Let (L_j, U_j) be a $100(1 - \alpha)\%$ confidence interval for θ_j and $S(\cdot)$ be any increasing function, then $\{S(L_j), S(U_j)\}$ is a $100(1 - \alpha)\%$ confidence interval for $S(\cdot)$, and if $S(\cdot)$ be any decreasing function then $\{S(U_j), S(L_j)\}$ is a $100(1 - \alpha)\%$ confidence interval for $S(\theta_j)$. Therefor, using (14), one may construct a confidence interval for other parameters of interest in the paper.

4 Illustrative example

To illustrate the proposed procedure in this paper, we consider a numerical example. Assuming $w_1 = 15$, a random ample of size 30 has been generated from ced in (2) with $p_1 = 0.015$ and $p_2 = 0.056$. The results are presented in Table 1.

Table 1. Generated sample of size 30 from ced in (2) with $w_1 = 15$, $p_1 = 0.015$ and $p_2 = 0.056$.

Parameter	Failure times										
p_1	4	7	9	13	15						
p_2	16	17	18	18	19	19	20	20	21	23	24	25
	26	28	29	31	32	35	36	40	40	43	48	58

Using the data in Table 1, we would obtain $N_1 = 5$ is the number of units that fail at stress level s_0. To investigate the variety of inference, we use different choices of w_2. From Table 1, for $w_2 = 20, 28$ and 35, the number of units that fail at stress level s_1 is given by $N_2 = 8, 14$ and 18, respectively.

The values of the point estimation for p_j, μ_j and $\bar{F}_j(9)$ $(j = 1, 2)$ have been obtained using (10), (11) and (12), respectively. Similar results for the interval estimation have been derived using (14). Toward this end, we consider the beta prior distributions with parameters (0.43, 23.4) and (2.17, 38.6) for p_1 and p_2, respectively. The results are summarized in Table 2.

Table 2. Values of Bayesian point and interval estimation for the parameters of interest based on the data in Table 1 for $w_1 = 15$ and some choices of w_2.

	$w_2 = 20$		$w_2 = 28$		$w_2 = 35$	
	point	interval	point	interval	point	interval
p_1	0.0122	(0.0042, 0.0242)				
p_2	0.0666	(0.0328, 0.1110)	0.0604	(0.0352, 0.0918)	0.0608	(0.0377, 0.0889)
μ_1	100.64	(41.310, 237.15)				
μ_2	16.5507	(9.0089, 30.441)	17.5854	(10.891, 28.437)	17.2546	(11.253, 26.504)
$\bar{F}_1(9)$	00.8859	(00.8021 0.9627)				
$\bar{F}_2(9)$	00.5117	(00.3468, 0.7404)	0.5417	(0.4203, 0.7246)	0.5384	(0.4328, 0.7074)

References

[1] Bagdonavicius, V. and Nikulin, M. (2002). Accelerated Life Models: Modeling and Statistical Analysis. Chapman & Hall/CRC Press, Boca Raton, FL.

[2] Balakrishnan, N., Cramer, E. and Dembińska, A. (2011). Characterizations of geometric distribution through progressively Type-II right-censored order statistics. *To appear in Statistics.*

[3] Balakrishnan, N. and Han, D. (2008). Exact inference for a simple step-stress model with competing risks for failure from exponential distribution under Type-II censoring. *J. Statist. Plann. Inference*, **138**, 4172–4186.

[4] Balakrishnan, N. and Xie, Q. (2007). Exact inference for a simple step-stress model with Type-I hybrid censored data from the exponential distribution. *J. Statist. Plann. Inference*, **137**, 3268–3290.

[5] Balakrishnan, N., Xie, Q. and Kundu, D. (2009). Exact inference for a simple step-stress model from the exponential distribution under time constraint. *Ann. Inst. Stat. Math.*, **61**, 251–274.

[6] Davarzani, N. and Parsian, A. (2011). Statistical Inference for Discrete Middle-Censored Data. *To appear in J. Statist. Plann. Inference.*

[7] DeGroot, M.H. and Goel, P.K. (1979). Bayesian estimation and optimal design partially accelerated life testing. *Naval Research Logistics Quarterly*, **26**, 223–235.

[8] Erto, P. and Giorgio, M. (2002). Assessing high reliability via Bayesian approach and accelerated tests. *Reliability Engineering System Safety*, **79**, 301–310.

[9] Gan, G. and Bain, L.J. (1995). Distribution of order statistics for discrete parents with applications to censored sampling. *J. Statist. Plann. Inference*, **44**, 37–46.

[10] Nagaraja, H.N. (1992). Order Statistics from Discrete Distributions. *Statistics*, **23**, 189–216.

[11] Nelson, W. (1990). Accelerated Testing: Statistical Models, Test Plans and Data Analysis. Wiley, New York.

[12] Rezaei, A.H. and Arghami, N.R. (2002). Right censoring in a discrete life model. *Metrika*, **55**, 151–160.

[13] Van Dorp, J.R., Mazzuchi, T.A., Fornell, G.E. and Pollock, L.R.A. (1996). Bayes approach to step-stress accelerated life testing. *IEEE Transactions on Reliability*, **45(3)**, 491–498.

[14] Van Dorp, J.R. and Mazzuchi, T.A. (2004). A general Bayes exponential inference model for accelerated life testing. *J. Statist. Plann. Inference*, **119**, 55–74.

[15] Van Dorp, J.R. and Mazzuchi, T.A. (2005). A general Bayes Weibull inference model for accelerated life testing. *Reliability Engineering System Safety*, **90**, 140–147.